

0420
#7300

OIKE

RAW SEQUENCE LISTING

DATE: 12/26/2001

PATENT APPLICATION: US/09/880,708B

TIME: 15:33:40

Input Set : N:\Crif3\RULE60\09880708B.txt

Output Set: N:\CRF3\12262001\I880708B.raw

SEQUENCE LISTING

3 (1) GENERAL INFORMATION:
5 (i) APPLICANT: Lee, Se-Jin
6 Huynh, Thanh
8 (ii) TITLE OF INVENTION: GROWTH DIFFERENTIATION FACTOR-5
10 (iii) NUMBER OF SEQUENCES: 27
12 (iv) CORRESPONDENCE ADDRESS:
13 (A) ADDRESSEE: Fish & Richardson, P.C.
14 (B) STREET: 4225 Executive Square, Suite 1400
15 (C) CITY: La Jolla
16 (D) STATE: CA
17 (E) COUNTRY: USA
18 (F) ZIP: 92037
20 (v) COMPUTER READABLE FORM:
21 (A) MEDIUM TYPE: Diskette
22 (B) COMPUTER: IBM Compatible
23 (C) OPERATING SYSTEM: Windows95
24 (D) SOFTWARE: FastSEQ for Windows Version 2.0
26 (vi) CURRENT APPLICATION DATA:
C--> 27 (A) APPLICATION NUMBER: US/09/880,708B
C--> 28 (B) FILING DATE: 12-Jun-2001
30 (vii) PRIOR APPLICATION DATA:
32 (A) APPLICATION NUMBER: 09/145,060
33 (B) FILING DATE:
35 (A) APPLICATION NUMBER: 08/455,559
36 (B) FILING DATE: 31-MAY-1995
38 (A) APPLICATION NUMBER: 08/003,144
39 (B) FILING DATE: 12-JAN-1993
41 (viii) ATTORNEY/AGENT INFORMATION:
42 (A) NAME: Lisa A. Haile, Ph.D.
43 (B) REGISTRATION NUMBER: 38,347
44 (C) REFERENCE/DOCKET NUMBER: 07265/057001
46 (ix) TELECOMMUNICATION INFORMATION:
47 (A) TELEPHONE: 619/678-5070
48 (B) TELEFAX: 619/678-5099
51 (2) INFORMATION FOR SEQ ID NO: 1:
53 (i) SEQUENCE CHARACTERISTICS:
54 (A) LENGTH: 28 base pairs
55 (B) TYPE: nucleic acid
56 (C) STRANDEDNESS: single
57 (D) TOPOLOGY: linear
59 (ii) MOLECULE TYPE: Genomic DNA
61 (vii) IMMEDIATE SOURCE:
62 (B) CLONE: 136
64 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:
66 CCGGAATTCG GNTGGGARMG NTGGRTNR
69 (2) INFORMATION FOR SEQ ID NO: 2:

ENTERED

28

RAW SEQUENCE LISTING

DATE: 12/26/2001

PATENT APPLICATION: US/09/880,708B

TIME: 15:33:40

Input Set : N:\Crf3\RULE60\09880708B.txt

Output Set: N:\CRF3\12262001\I880708B.raw

```

71      (i) SEQUENCE CHARACTERISTICS:
72          (A) LENGTH: 42 base pairs
73          (B) TYPE: nucleic acid
74          (C) STRANDEDNESS: single
75          (D) TOPOLOGY: linear
77      (ii) MOLECULE TYPE: Genomic DNA
79      (vii) IMMEDIATE SOURCE:
80          (B) CLONE: 121
82      (ix) FEATURE:
83          (A) NAME/KEY: Coding Sequence
84          (B) LOCATION: 1..42
C--> 85          (D) OTHER INFORMATION: / N at residue 13, 25 and 28 = Inosine
87      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
89      CCGGAATTCR CANCCRCAYT CRTCNACNAC CATRTCCTYTCR TA      42
92      (2) INFORMATION FOR SEQ ID NO: 3:
94          (i) SEQUENCE CHARACTERISTICS:
95              (A) LENGTH: 7 amino acids
96              (B) TYPE: amino acid
97              (D) TOPOLOGY: linear
99          (ii) MOLECULE TYPE: peptide
101         (vii) IMMEDIATE SOURCE:
102             (B) CLONE: 136
104         (ix) FEATURE:
105             (B) LOCATION: 1..7
C--> 106         (D) OTHER INFORMATION: Xaa at residue 4 = Arg or Ser; Xaa at residue
C--> 107 6 and 7 = Val, Ile or Met
109         (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:
W--> 111 Gly Trp Glu Xaa Trp Xaa Xaa
112         1          5
115      (2) INFORMATION FOR SEQ ID NO: 4:
117          (i) SEQUENCE CHARACTERISTICS:
118              (A) LENGTH: 11 amino acids
119              (B) TYPE: amino acid
120              (D) TOPOLOGY: linear
122          (ii) MOLECULE TYPE: peptide
124          (vii) IMMEDIATE SOURCE:
125              (B) CLONE: 121
127          (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:
129      Tyr Glu Asp Met Val Val Asp Glu Cys Gly Cys
130      1          5          10
133      (2) INFORMATION FOR SEQ ID NO: 5:
135          (i) SEQUENCE CHARACTERISTICS:
136              (A) LENGTH: 35 base pairs
137              (B) TYPE: nucleic acid
138              (C) STRANDEDNESS: single
139              (D) TOPOLOGY: linear
141          (ii) MOLECULE TYPE: Genomic DNA
143          (vii) IMMEDIATE SOURCE:
144              (B) CLONE: 141

```

RAW SEQUENCE LISTING

DATE: 12/26/2001

PATENT APPLICATION: US/09/880,708B

TIME: 15:33:40

Input Set : N:\Crf3\RULE60\09880708B.txt

Output Set: N:\CRF3\12262001\I880708B.raw

```

146      (ix) FEATURE:
147          (A) NAME/KEY: Coding Sequence
148          (B) LOCATION: 1..35
C--> 149      (D) OTHER INFORMATION: N at residue 12, 27, 30 and 33 = Inosine
151      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:
153      CCGGAATTCG GNTGGVANRA YTGGRTNRTN KCNCC                      35
156      (2) INFORMATION FOR SEQ ID NO: 6:
158          (i) SEQUENCE CHARACTERISTICS:
159              (A) LENGTH: 33 base pairs
160              (B) TYPE: nucleic acid
161              (C) STRANDEDNESS: single
162              (D) TOPOLOGY: linear
164          (ii) MOLECULE TYPE: Genomic DNA
166      (vii) IMMEDIATE SOURCE:
167          (B) CLONE: 145
169      (ix) FEATURE:
170          (A) NAME/KEY: Coding Sequence
171          (B) LOCATION: 1..33
C--> 172      (D) OTHER INFORMATION: N at residue 13, 19, 25 and 28 = Inosine
175      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:
177      CCGGAATTCR CANSCRCANG MNTCNACNRY CAT                      33
180      (2) INFORMATION FOR SEQ ID NO: 7:
182          (i) SEQUENCE CHARACTERISTICS:
183              (A) LENGTH: 9 amino acids
184              (B) TYPE: amino acid
185              (D) TOPOLOGY: linear
187          (ii) MOLECULE TYPE: peptide
189      (vii) IMMEDIATE SOURCE:
190          (B) CLONE: 141
192      (ix) FEATURE:
193          (B) LOCATION: 1..9
C--> 194      (D) OTHER INFORMATION: Xaa at residue 3 = His, Gln, Asn, Lys, Glu or
C--> 195 Asp; Xaa at residue 4 = Asp or Asn; Xaa at residues 6 and 7 = Val, Ile or Met;
C--> 196 Xaa at residue 8 = Glu or Ser
198      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:
W--> 200      Gly Trp Xaa Xaa Trp Xaa Xaa Xaa Pro
201          1          5
204      (2) INFORMATION FOR SEQ ID NO: 8:
206          (i) SEQUENCE CHARACTERISTICS:
207              (A) LENGTH: 8 amino acids
208              (B) TYPE: amino acid
209              (D) TOPOLOGY: linear
211          (ii) MOLECULE TYPE: peptide
213      (vii) IMMEDIATE SOURCE:
214          (B) CLONE: 145
216      (ix) FEATURE:
217          (B) LOCATION: 1..8
C--> 218      (D) OTHER INFORMATION: Xaa at residues 2 and 3 = Val, Ile, Met, Thr
C--> 219 or Ala; Xaa at residue 4 = Asp or Glu; Xaa at residue 5 = Ala or Ser; Xaa at

```

RAW SEQUENCE LISTING

DATE: 12/26/2001

PATENT APPLICATION: US/09/880,708B

TIME: 15:33:40

Input Set : N:\Crif3\RULE60\09880708B.txt

Output Set: N:\CRF3\12262001\I880708B.raw

C--> 220 residue 7 = Gly, Ala, Arg, Asn, Asp, Cys, Glu, Gln, His, Ile, Leu, Lys, Met,

C--> 221 Phe Pro, Ser, Thr, Trp, Tyr, and Val.

223 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

W--> 225 Met Xaa Xaa Xaa Xaa Cys Xaa Cys

226 1 5

228 (2) INFORMATION FOR SEQ ID NO: 9:

230 (i) SEQUENCE CHARACTERISTICS:

231 (A) LENGTH: 2329 base pairs

232 (B) TYPE: nucleic acid

233 (C) STRANDEDNESS: single

234 (D) TOPOLOGY: linear

236 (ii) MOLECULE TYPE: Genomic DNA

238 (vii) IMMEDIATE SOURCE:

239 (B) CLONE: GD-5

241 (ix) FEATURE:

242 (A) NAME/KEY: Coding Sequence

243 (B) LOCATION: 322...1806

245 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

247	TTCAAGCCCT	CAGTCAGTTG	TGCGGGAGAA	AGGGGGCGGT	CGGCTTTCTC	CTTTCAAGAA	60
248	CGAGTTATTT	TCAGCTGCTG	ACTGGAGACG	GTGCACGTCT	GGACACGGGA	GCACTTCCAC	120
249	TATGGGACTG	GATACAGACA	CACGCCCGGC	GGACTTCAAG	ACACTCAGAC	TGAGGAGAAA	180
250	GCCCTGCCTG	CTGCTGCTGC	TGCTGCTGCT	GCCACCGCTG	CCTCTGAAGA	CCCACTCCTT	240
251	TCATGGTTTT	TCCTGCCAAG	CCAGAGGCAC	CTTCGCTGCT	ACGGCCTTTC	TCTGTGGTGT	300
252	CATTCAGCGG	CTGGCCAGAG	G ATG AGA CTC CCC AAA	CTC CTC ACT CTT TTG			351
253			Met Arg Leu Pro Lys	Leu Leu Thr Leu Leu			
254			1 5			10	
256	CTG TGG CAC CTG GCT TGG CTG GAC CTG GAA CTC ATC TGC ACT GTG CTG						399
257	Leu Trp His Leu Ala Trp Leu Asp Leu Glu Leu Ile Cys Thr Val Leu						
258		15		20		25	
260	GGT GCC CCT GAC TTA GGA CAG AGA ACC CCA GGG GCC AAG CCA GGG TTG						447
261	Gly Ala Pro Asp Leu Gly Gln Arg Thr Pro Gly Ala Lys Pro Gly Leu						
262		30		35		40	
264	ACC AAA GCG GAG GCC AAG GAG AGG CCA CCC CTG GCC AGG AAT GTC TTT						495
265	Thr Lys Ala Glu Ala Lys Glu Arg Pro Pro Leu Ala Arg Asn Val Phe						
266		45		50		55	
268	AGG CCA GGG GGT CAT ATC TAT GGT GTG GGG GCC ACC AAT GCC AGG GCC						543
269	Arg Pro Gly Gly His Ile Tyr Gly Val Gly Ala Thr Asn Ala Arg Ala						
270		60		65		70	
272	AAG GGA AGC TCT GGG CAG ACA CAG GCC AAG AAG GAT GAA CCC AGA AAG						591
273	Lys Gly Ser Ser Gly Gln Thr Gln Ala Lys Lys Asp Glu Pro Arg Lys						
274		75		80		85	90
276	ATG CCC CCC AGA TCC GGT GGC TCT GAA ACC AAG CCA GGA CCC TCT TCC						639
277	Met Pro Pro Arg Ser Gly Gly Ser Glu Thr Lys Pro Gly Pro Ser Ser						
278		95		100		105	
280	CAG ACT AGA CAG GCT GCA GCC CGG ACT GTA ACC CCA AAA GGA CAG CTT						687
281	Gln Thr Arg Gln Ala Ala Ala Arg Thr Val Thr Pro Lys Gly Gln Leu						
282		110		115		120	
284	CCT GGG GGC AAA GCA TCT TCA AAA GCA GGA TCT GCC CCC AGC TCC TTC						735
285	Pro Gly Gly Lys Ala Ser Ser Lys Ala Gly Ser Ala Pro Ser Ser Phe						

DATE: 12/26/2001

TIME: 15:33:40

Output Set: N:\CRF3\12262001\I880708B.raw

file:///C:/CRF3/Outhold/VsrI880708B.htm

VERIFICATION SUMMARY

DATE: 12/26/2001

PATENT APPLICATION: US/09/880,708B

TIME: 15:33:41

Input Set : N:\Crf3\RULE60\09880708B.txt

Output Set: N:\CRF3\12262001\I880708B.raw

L:27 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]
L:28 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]
L:85 M:220 C: Keyword misspelled or invalid format, [(D) OTHER INFORMATION:]
L:106 M:220 C: Keyword misspelled or invalid format, [(D) OTHER INFORMATION:]
L:111 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3
L:149 M:220 C: Keyword misspelled or invalid format, [(D) OTHER INFORMATION:]
L:172 M:220 C: Keyword misspelled or invalid format, [(D) OTHER INFORMATION:]
L:194 M:220 C: Keyword misspelled or invalid format, [(D) OTHER INFORMATION:]
L:200 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:7
L:218 M:220 C: Keyword misspelled or invalid format, [(D) OTHER INFORMATION:]
L:225 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8